



**HAL**  
open science

## Writing e-Exams in Pre-University College

Mathew Hillier, Nathaniel Lyon

► **To cite this version:**

Mathew Hillier, Nathaniel Lyon. Writing e-Exams in Pre-University College. Open Conference on Computers in Education (OCCE), Jun 2018, Linz, Austria. pp.264-274, 10.1007/978-3-030-23513-0\_26 . hal-02370926

**HAL Id: hal-02370926**

**<https://inria.hal.science/hal-02370926>**

Submitted on 19 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Writing e-Exams in Pre-University College

Mathew Hillier

Monash University, Australia, Mathew.Hillier@monash.edu

Nathaniel Lyon

Monash College, Australia, nathaniel.lyon@monashcollege.edu.au

**Abstract.** This study examined students' expressed strategies, habits and preferences with respect to responding to supervised text-based assessments. Two trials of a computerised examination system took place in an Australian pre-university college in 2016 and 2017. Students in several classes studying geography and globalisation completed a sequence of practice and assessed work. Data were collected using pre- and post-surveys about their preferred writing styles, habits and strategies in light of their choice to type or handwrite essay and short answer examinations. Comparisons were made between those that elected to handwrite and those who chose to type the examination, with several areas being significant. The performance (grades), production (word count) of the typists and hand-writers were also correlated and compared.

**Keywords.** e-Exams, writing strategies, student perceptions, affordances

## 1 Introduction and background

This study is part of a nationally funded project [1 - 3] looking at modernising supervised high stakes assessment within the Australian higher education context. This paper builds on previous e-examination (e-Exam or e-exam) trials held in Australian universities, investigating how students in a pre-tertiary pathway college context perceive the task of writing e-Exams.

We use the term e-Exam to specifically refer to a “timed, supervised, summative assessment conducted using each candidate’s own computer running a standardised operating system” [4]. This differentiates our work from those employing ‘online’ testing tools that take the format of a data collection instrument (e.g. Moodle quiz, Blackboard test, TCExam, QuestionMark Perception, ExamSoft). In our case, we include the use of ‘authentic’ software applications fit for the purpose of the assessment task, e.g. a full office suite is provided to write reports or essays.

We will explore the literature related to the writing of high stakes assessment tasks and in particular computerised examinations before moving on to explain the process used in the study. This is followed by survey results and a discussion of the findings that draws out implications for future research and practice.

## 2 Literature

The idea of using computers for assessment has been around for 60 years [5], yet pen-on-paper still dominates most higher education and school examinations around the

world. However, attention has recently shifted to modernising the examination room. The ‘Dublin Declaration’ [6] features the idea of ‘authentic assessment’ [7] as one of its core recommendations when considering the use of computers for assessment. The Declaration specifically recommends that an e-Exam must be an:

“authentic assessment that matches modern workplace practices and many student learning experiences”. (p. xviii)

If we consider the modern classroom, or the majority of work places, we find that computer and software technology is near ubiquitous. From accountancy to zoology, computers are now key tools of the trade. From keeping track of laboratory experiments to report writing and bookkeeping, we could be confident that a computer was involved. Similarly, in higher education, reports, essays and communications are conducted via the typed medium. The vast majority of students in the developed world today use software tools as part of their course work and in addressing unsupervised assessment tasks. University students could hardly remember or even know what it would be like to hand-write their class work, yet we ask them to write examinations that can take up to three or more hours. The keyboard is now so commonly used that the very skill of handwriting is in decline [8], with a subsequent loss of the motor skills required to write proficiently. A previous study [9] found that hand-writers in examinations beyond about 70 minutes felt physical discomfort while typists were unaffected. Fortunately, there are a number of examples where the transition to e-exams is well underway or at least beginning. Prior work [10, 4] described and reviewed several e-Exam projects, with a longer list available [11].

As part of the conversion effort, teachers and students may have an adjustment period when transitioning from hand-writing under examination conditions to typing. A survey [12] of students at the University of Bradford in the United Kingdom (UK) on their experience of using QuestionMark Perception covered a range of topics. Differences between handwriting and typing an essay-style examination were explored [13], and the authors looked at self-reported typing prowess, confidence, stress, use of time, writing strategies, pre-planning, structuring, editing and reviewing prior to submission. Similarly work by other authors [14] looked at the preferences of students with regard to typing or handwriting essay examinations when they were given the choice. They noted a 10% uptake by students of the typing option. Another researcher [15] reported on students’ preconceptions of what an e-exam may involve prior to the start of several e-Exam trials at an Australian university, with interest by students varying significantly between discipline areas. A follow up study of e-exam trials in six courses at the same institution [9] showed the typing option was selected by 5% to 34% of students. The top three comments from typists were that they could type faster, it was neater than handwriting, and that they were easily able to edit their work leading to more polished responses.

Finally, we turn to the matter of performance. Stakeholders are concerned that computerising the examination room may have an uneven impact on student performance by advantaging some and disadvantaging others [16]. Concerns over typing prowess, speed, and computer access have been raised. However, similar concerns could be raised with respect to the uneven abilities of students in terms of speed, neatness and physical strength for handwriting long examinations (ibid). Performance in an e-exam in a dentistry unit found that there was a moderate increase in marks for those that typed the examination compared to those handwriting. There

could be no causal link established, due to the self-selecting sample groups [2]. Hand-writers that produced more words generally did better while more typed words did not see marks increase by as much. Another study [17] compared typists and hand-writers and found that the typists produced around 20% more words in an e-exam. Similarly, a further study [16] found that typists produced more words than hand-writers but there was no statistically significant difference in grades. In terms of presentation effects, one researcher [18] found no significant difference in scores due to presentation mode (typed or transcribed handwritten) on a large-scale writing assessment. Prior studies have also looked at demographics such as gender [19] and how this may impact a student's performance when faced with a typed examination.

The review of prior work has outlined several areas of focus that informed questions to be investigated (see Table 1), and provided an up-front frame for analysis.

**Table 1.** Areas for investigation

| <b>Area</b>           | <b>Example research questions</b>   |
|-----------------------|---|
| Rationale of students | What proportion of candidates were in favour of typing their examination? What rationale was provided for their choice? Did the e-Exam environment support their writing? |
| Writing strategies    | Were there differences in the writing preferences and strategies used by those that typed and those that handwrote?   |
| Student performance   | Were there differences in words produced and grades achieved by those that typed and those that handwrote?  |

### 3 Study context

In Australia, the first author is leading an Australian federal government funded grant project [1] investigating how authentic e-assessment can be introduced into the examination rooms of universities using bring-your-own-devices (BYOD). This study was undertaken within two units offered in the Foundation Year [20] programme at Monash College, Australia. The programme is the equivalent of an Australian year 12 high school certificate or the International Baccalaureate. The second author is a unit coordinator and teacher of the units in which the e-exam trials were conducted. The trials were carried out using in-class supervised written assessments requiring short text and essay-style responses.

### 4 Method and approach

Students in several classes undertaking geography (Geo) in semester 1, 2016 and globalisation (Glo) in semester 2, 2017 at Monash College took part in a two phase trial of the e-Exam system. Ethics protocol approval was gained via Monash University prior to the trials. In this study a fully functional word processor was used that was part of the e-Exam platform as previously described [3]. Students used a custom Live Linux universal serial bus (USB) environment on their own laptops that provided a full office

suite. They could use editing tools such as spelling, grammar and highlighting to help with their writing tasks. Being able to copy, delete or move text around with ease meant that there were functional differences to undertaking the same task using pen-on-paper.

In phase one, students participated in an in-class, ungraded preview session. They were provided with instructions on how to use the e-Exam system with laptops and opportunity to practice e-exam processes, use the software, and try the question response format. Students completed a pre-examination survey requesting technical information about their laptop and first impressions of the e-Exam process and software. Attitudes and responses to e-Exam software use are reported separately [21].

Phase two occurred two weeks later in the form of an in-class, graded, supervised assessment task; students could choose to type or handwrite. Materials were provided on paper or as a word processor document and included photographs, diagrams, charts and data tables. In Geo, the assessment task was a single case study with an extended essay response. In Glo, two short answer sections and a mini-case essay response were required. Students then completed a post-examination survey before leaving the room. An extended account of the trial process is provided elsewhere [21].

Qualitative survey data relating to students' opinions on writing in examinations was analysed using SPSS v24. Likert items were treated as non-parametric as per advice [22]. This stance is supported by other authors [12] when analysing students' perceptions of e-Assessment. The Mann-Whitney U test [23] was used to test the variance between groups (males versus females and typists versus hand-writers) of Likert responses. Chi-squared was also used to test if gender played a role in the decision to type the examination. A Fisher exact test [23] was used when comparing categorical variables. When it came to performance data (grades) a T-test was used to compare between groups while a Spearman test [25] was used to test for correlation between word count and grades. We used an alpha level of .05 for all statistical tests, unless otherwise noted. However, the participants were not randomly assigned to the typing or hand-writing groups, which makes the results only descriptive of this study. In terms of items reflecting students' opinions, we used statistical tests as a tool to summarise rather than to be representative of an objective truth [see 12].

## 5 Findings

The trials involved 128 pre-tertiary students; 65% female and 35% male. In the examination, 52% of students elected to type. Table 2 shows participation at each stage.

**Table 2.** Intention to type the examination at each stage of the study (counts)

| <b>Participation</b> | <b>Initial interest</b> | <b>After practice</b> | <b>At the examination</b> |
|----------------------|-------------------------|-----------------------|---------------------------|
| Type                 | 73                      | 72                    | 64                        |
| Handwrite            | 26                      | 30                    | 59                        |
| Missing              | 29                      | 26                    | 5                         |

A Chi-squared test revealed a statistically significant difference in relation to gender and choice to type the examination ( $\chi^2(1)=5.299, p=0.021$ ) with 68% of males choosing to type compared to 46% of females. However, the differences due to gender were not

significant in the earlier stages of the trial. This may be due to the larger number of undecided ('missing') cases in the earlier stages.

### Opinions

We examined the differences in opinions on five-point Likert items (where strongly agree=5, neutral=3, and strongly disagree=1) between those that typed the examination and those that handwrote it using a Mann-Whitney U test. Table 3 displays the extent of agreement, with means and standard deviations provided for clarity, along with the results of significance tests with the difference between the means shown.

**Table 3.** Post-examination survey responses by text production mode

| Question  | Typed    |          |           | Handwrote |          |           | MW   |          |          |
|---|----------|----------|-----------|-----------|----------|-----------|------|----------|----------|
|   | <i>n</i> | <i>M</i> | <i>SD</i> | <i>N</i>  | <i>M</i> | <i>SD</i> | diff | <i>U</i> | <i>p</i> |
| <b>Post-examination Survey</b>                                    |          |          |           |           |          |           |      |          |          |
| I type faster than I handwrite                                    | 64       | 4.1      | 1.1       | 55        | 1.8      | 1.2       | 2.2  | 412      | >.001    |
| I type accurately   | 63       | 4.1      | 0.8       | 52        | 2.5      | 1.1       | 1.7  | 379      | >.001    |
| When I make errors, I am able to quickly correct them when typing | 64       | 4.3      | 0.8       | 53        | 2.9      | 1.3       | 1.4  | 634      | >.001    |
| I often rely on spell check to detect errors                      | 62       | 3.8      | 1.3       | 54        | 2.9      | 1.3       | 0.9  | 1060.5   | >.001    |
| I work more efficiently when I type on a familiar keyboard        | 64       | 4.2      | 0.9       | 53        | 3.0      | 1.3       | 1.2  | 721.5    | >.001    |
| My hand-writing is normally neat and legible                      | 63       | 3.6      | 1.3       | 52        | 3.1      | 1.1       | 0.5  | 1183.5   | 0.008    |
| I go back to re-read and revise my writing quite a lot            | 63       | 3.8      | 1.0       | 53        | 2.9      | 1.2       | 0.9  | 969.5    | >.001    |
| I prepare most of my assignments /reports using a computer        | 64       | 4.0      | 0.9       | 53        | 3.0      | 1.4       | 1.0  | 1012.5   | >.001    |

Those that chose to hand-write the examination were asked about the neatness of their writing and comfort levels experienced during the examination (see Table 4). The examination duration was 70 minutes in both cases.

**Table 4.** Post-examination survey responses on hand-writing

| Question   | <i>n</i> | <i>M</i> | <i>SD</i> |
|--|----------|----------|-----------|
| I think my hand writing was neat and legible                   | 53       | 3.6      | 0.9       |
| I experienced discomfort (sore/tired/cramp) in my writing hand | 53       | 2.9      | 1.2       |

Typists in the geography unit were also asked to reflect specifically on using a computer for the examination, given the nature of the task (see Table 5).

**Table 5.** Post-examination survey responses on using a computer for the assessment

| Question   | <i>n</i> | <i>M</i> | <i>SD</i> |
|--|----------|----------|-----------|
| I was able to produce a better final version of this assessment      | 24       | 4.3      | 0.6       |
| I was able to quickly complete the assessment                        | 24       | 4.3      | 0.7       |
| I was able to easily edit and make changes                           | 24       | 4.6      | 0.6       |
| I was able to easily refer to reference materials and resources      | 24       | 3.7      | 1.0       |
| I was able to easily think and compose my answer using a computer    | 24       | 4.3      | 0.6       |
| I would like to use a computer for similar assessments in the future | 23       | 4.2      | 0.7       |

Students were asked about their preferences for production method ('computer' (C), 'same' (S) or 'pen and paper' (P)) for a range of writing activities, style and features. A Fisher exact test for categorical variables was used to see if their preferences may have influenced their choice to type or handwrite the examination (see Table 6). Counts and percentages in brackets are shown for each production method (C, S or P) for the two groups (typists and hand-writers). Significant differences were noted across many of the items with preferences in alignment with their actual choice of examination mode.

**Table 6.** Post-examination survey responses on examination writing strategies

| Question  | Typed      |            |            |            | Handwrote  |            |            |            | Fisher<br><i>p</i> |
|---|------------|------------|------------|------------|------------|------------|------------|------------|--------------------|
|   | C          | S          | P          | <i>n</i>   | C          | S          | P          | <i>n</i>   |                    |
| C=computer, S=same, P=pen   |            |            |            |            |            |            |            |            |                    |
| I write more words in an examination when...                              | 35<br>(56) | 23<br>(37) | 5<br>(8)   | 63<br>(53) | 1<br>(2)   | 18<br>(32) | 37<br>(66) | 56<br>(47) | >.001              |
| I write faster in an examination when...                                  | 42<br>(66) | 17<br>(27) | 5<br>(8)   | 64<br>(53) | 3<br>(5)   | 9<br>(16)  | 44<br>(79) | 56<br>(47) | >.001              |
| I think more carefully before I start writing in an examination when...   | 32<br>(50) | 13<br>(20) | 19<br>(30) | 64<br>(54) | 3<br>(5)   | 17<br>(31) | 35<br>(64) | 55<br>(46) | >.001              |
| I pause to think most in an examination when...                           | 21<br>(33) | 19<br>(30) | 24<br>(38) | 64<br>(53) | 11<br>(20) | 26<br>(46) | 19<br>(34) | 56<br>(47) | 0.123              |
| I write in a style that feels more normal in an examination when...       | 29<br>(46) | 24<br>(38) | 10<br>(16) | 63<br>(53) | 1<br>(2)   | 17<br>(30) | 38<br>(68) | 56<br>(47) | >.001              |
| I try not to make changes unless they are really important when...        | 18<br>(28) | 18<br>(28) | 28<br>(44) | 64<br>(53) | 3<br>(5)   | 25<br>(45) | 28<br>(50) | 56<br>(47) | 0.003              |
| I change, move or correct words or phrases most when...                   | 42<br>(67) | 13<br>(21) | 8<br>(13)  | 63<br>(53) | 11<br>(20) | 21<br>(38) | 24<br>(43) | 56<br>(47) | >.001              |
| I think the overall structure/argument of my response is better when...   | 38<br>(59) | 15<br>(23) | 11<br>(17) | 64<br>(53) | 3<br>(5)   | 21<br>(38) | 32<br>(57) | 56<br>(47) | >.001              |
| I make more effective use of the time available in an examination when... | 43<br>(67) | 16<br>(25) | 5<br>(8)   | 64<br>(53) | 2<br>(4)   | 15<br>(27) | 39<br>(70) | 56<br>(47) | >.001              |
| I go back and read over my response most in an examination when...        | 28<br>(44) | 28<br>(44) | 8<br>(13)  | 64<br>(53) | 3<br>(5)   | 23<br>(41) | 30<br>(54) | 56<br>(47) | >.001              |
| I feel more stressed in an examination when...                            | 15<br>(23) | 28<br>(44) | 21<br>(33) | 64<br>(53) | 18<br>(32) | 22<br>(39) | 16<br>(29) | 56<br>(47) | 0.589              |
| I am more likely to run out of time in an examination when...             | 13<br>(20) | 28<br>(44) | 23<br>(36) | 64<br>(53) | 16<br>(29) | 22<br>(39) | 18<br>(32) | 56<br>(47) | 0.601              |
| Overall, I feel I perform better in an examination when...                | 37<br>(59) | 22<br>(35) | 4<br>(6)   | 63<br>(53) | 1<br>(2)   | 14<br>(25) | 41<br>(73) | 56<br>(47) | >.001              |

### Performance

Performance was expressed as a percentage grade. There was a statistically significant performance difference between the two units (Geo,  $n=38$ ,  $Mdn$  73.75 and Glo,  $n=85$ ,  $Mdn$  57.14) using a Mann-Whitney U test ( $U=1100$ ,  $p=.005$ ).

When the grade data were grouped by gender (shown in Table 7) it was not found to be statistically significant, although the result was borderline at  $p=.050$ . In the geography unit, a significant difference in grades between genders was found ( $p=.006$ );

however, this was not so in globalisation. Normal distributions were confirmed for each gender pair using standardised skewness and the Shapiro-Wilks test with Levene's test (shown in Table 7) demonstrating equivalence of variance. Means and standard deviations are also provided for clarity.

**Table 7.** Grades out of 100 by gender

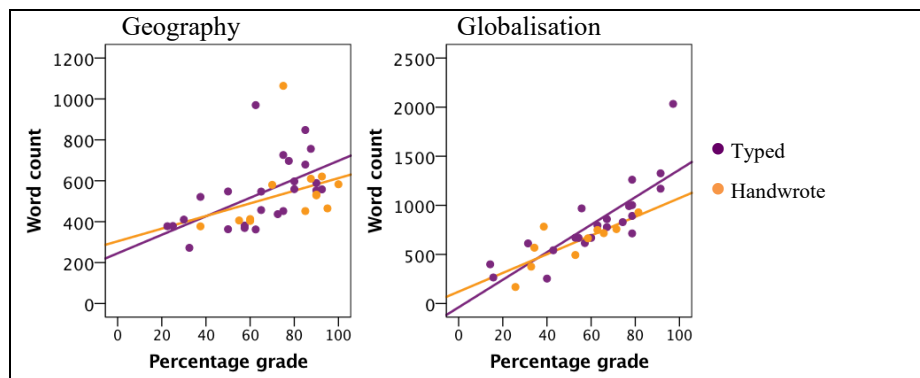
| Unit | Female |      |      | Male |      |      | Df      | T-test |              | Levene's test |       |
|------|--------|------|------|------|------|------|---------|--------|--------------|---------------|-------|
|      | n      | M    | SD   | n    | M    | SD   |         | T      | P            | F             | p     |
| Geo  | 24     | 76.2 | 17   | 12   | 56.4 | 23   | (1,34)  | 2.912  | <b>0.006</b> | 1.422         | 0.241 |
| Glo  | 53     | 58.3 | 18.5 | 29   | 56   | 20.7 | (1,80)  | 0.519  | 0.605        | 0.002         | 0.961 |
| Both | 77     | 63.9 | 19.8 | 41   | 56.1 | 21.1 | (1,116) | 1.98   | 0.050        | 0.005         | 0.943 |

When comparing the grades of typists and hand-writers (shown in Table 8) across both units, no statistically significant difference was found. However, within the globalisation unit there was a small statistically significant difference in grades between typists and hand-writers ( $p=.033$ ). The distribution of each pair was found to be normal using standardised skewness and the Shapiro-Wilks tests, with Levene's test (shown in Table 8) establishing equivalence of variance. Means and standard deviations are provided for clarity.

**Table 8.** Grades out of 100 for typists and hand-writers

| Unit | Typed |      |      | Handwrote |      |      | Df      | T-test |              | Levene's test |       |
|------|-------|------|------|-----------|------|------|---------|--------|--------------|---------------|-------|
|      | n     | M    | SD   | n         | M    | SD   |         | T      | p            | F             | p     |
| Geo  | 25    | 64.3 | 21.5 | 13        | 76.7 | 18.8 | (1,36)  | -1.754 | 0.088        | 0.212         | 0.648 |
| Glo  | 39    | 62.9 | 21.0 | 44        | 53.7 | 17.4 | (1,81)  | 2.174  | <b>0.033</b> | 1.636         | 0.205 |
| Both | 64    | 63.4 | 21.0 | 57        | 59.0 | 20.1 | (1,119) | 1.190  | 0.236        | 0.333         | 0.565 |

Within each of the geography group ( $n=38$ ) and the globalisation group ( $n=32$ ) positive, statistically significant correlations were found between the number of words written and the grade achieved from Spearman tests (Geo:  $r_s=.628$ ,  $p=>.001$  and Glo:  $r_s=.865$ ,  $p=>.001$ ). See Figure 1.



**Figure 1.** Correlation between number of words and grade



Mann-Whitney U tests found no statistically significant difference in the number of words produced by typists and hand-writers (see Table 9). It is worth noting that word count data were only available for half the globalisation group and that assessment tasks were different between Glo and Geo, so the two groups are not comparable.

**Table 9.** Words produced by typists and hand-writers

| Unit          | Typed    |          |           | Handwrote |          |           | diff | MW       |          |
|---------------|----------|----------|-----------|-----------|----------|-----------|------|----------|----------|
|               | <i>n</i> | <i>M</i> | <i>SD</i> | <i>n</i>  | <i>M</i> | <i>SD</i> |      | <i>U</i> | <i>p</i> |
| Geography     | 25       | 536      | 170       | 13        | 541      | 178       | 5    | 157.5    | 0.878    |
| Globalisation | 22       | 833      | 83        | 10        | 620      | 71        | 213  | 71       | 0.113    |

## 6 Discussion

The choice to type or handwrite appears to be multifaceted, but with the strongest differences declared around being able to type more quickly than handwrite and typing accuracy. Overall, the results showed an alignment between writing strategies and choice of text production mode in the examination (see Tables 3 and 6). The result of 52% of students electing to type is much higher than reported in previous studies involving free choice [9, 14]. The use of in-class practice sessions compared to optional out-of-class practice sessions used [9] may have contributed to the increased uptake of the typing option in our study.

In both 70-minute assessment tasks, only a minority of hand-writers experienced discomfort (i.e. agreement of 2.9 out of a maximum 5 with standard deviation 1.2 indicating that only some were impacted). This was consistent with previous work [9] where a 70-minute duration was the cross-over point where discomfort started to become a problem.

Typists generally agreed that the computer allowed them to produce better responses on the assessment task (i.e. most Likert items in Table 5 were rated above 4 on the 5-point scale), including that they would like to be able to use a computer for similar assessments in the future. The weakest agreement related to their ability to easily refer to reference materials. In this study, students had to scroll up and down between the materials and their written response. Providing reference materials in a separate file (that would allow side-by-side window arrangement or ‘Alt-Tab’ between windows), in a split screen format within the document [26] or on paper, may help in this regard.

In terms of performance, there were mixed results. The number of words produced by typists was greater in Glo, as shown elsewhere [16, 17], but less in Geo. The classes were relatively small and assessment tasks were different, so this may have played a role. In Geo, hand-writers achieved slightly better grades (but not significantly so) and in Glo, typists did better with a borderline statistically significant difference in grades. When comparing by gender, in Geo, females did significantly better. In Glo, they also did slightly better, but not significantly so. Overall, a weak link between gender and performance could be claimed within these groups. In both classes, we found a statically significant correlation between words produced and marks. The scatter plot in Figure 1 shows that the correlation occurred over a narrow band in Geo, with a minimum of 300 words required before a pass was possible. Yet an adequate word

count certainly does not guarantee a pass. Similarly, a larger number of words above 600 did not see the highest grades; indeed, those with the highest grades wrote about the same as the group mean of 538 words. The scatter plot for Glo shows a more tightly monotonic relationship between words and grades. Again, correlation is certainly not causation in terms of word count – quality still matters!

## 7 Conclusion

Overall, we observed a significant alignment between preferred writing strategies and choice of text production method in the two trials. Further, the grades achieved between typists and hand-writers did not differ significantly. These two facts are not surprising when dealing with thinking, purposeful humans who aim to maximise outcomes. However, this does raise the concern that should a shift occur towards fully-typed assessment, then a period of transition with assistance for those who preferred handwriting should occur to ensure they are able to adapt successfully.

**Acknowledgments.** The authors would like to thank the Australian Government Department of Education and Training for financial support and the students at Monash College for being willing to ‘give it a go’.

## References

1. TEAA: Transforming exams across Australia. Australian Government Department of Education and Training, Grant ID15-4747 (2015). Retrieved from: <http://transformingexams.com>
2. Fluck A., Hillier, M.: Innovative assessment with eExams. Presented at the Australian Council for Computers in Education Conference, Brisbane, 29 September to 2 October (2016). Retrieved from: <http://conference.acce.edu.au/index.php/acce/acce2016/paper/view/34/27>
3. Hillier, M., Fluck, A.: Transforming Exams - How IT works for BYOD e-Exams. Presented at the Australasian Society for Computers in Learning in Tertiary Education Conference, 100-105. Toowoomba, Australia (2017). Retrieved from: <http://2017conference.ascilite.org/wp-content/uploads/2017/11/Concise-HILLIER.pdf>
4. Fluck, A., Hillier, M.: eExams: Strength in Diversity. In A. Tatnall, M. Webb (Eds.) *Tomorrow’s Learning: Involving Everyone*. Learning with and about Technologies and Computing, Cham, Switzerland: Springer International Publishing, 515, 409–417 (2017)
5. Swets, J.A., Feurzeig, W.: Computer-Aided Instruction. *Science*, 150(3696), 572–576 (1965)
6. Tatnall, A., Webb, M. (Eds.): *Tomorrow’s Learning: Involving Everyone*. Learning with and about Technologies and Computing. Heidelberg, Germany: Springer (2017)
7. Crisp, G.: Towards Authentic e-Assessment Tasks. Presented at the EdMedia: World Conference on Educational Media and Technology, Honolulu, HI, 1585–1590 (2009). Retrieved from: <http://www.editlib.org/p/31689/>
8. Sülzenbrück, S., Hegele, M., Rinkenauer, G., Heuer, H.: The Death of Handwriting: Secondary Effects of Frequent Computer Use on Basic Motor Skills. *Journal of Motivation and Behavior*, 43(3), 247–251 (2011)

9. Hillier, M.: To type or handwrite: student's experience across six e-Exam trials. In T. Reiners, B.R von Kinsky, D. Gibson, V. Chang, L. Irving, K. Clarke (eds.) Proceedings of the ASCILITE conference, Perth, Australia, 143-154 (2015). Retrieved from: [http://transformingexams.com/files/Hillier\\_2015\\_ascilite\\_fp.pdf](http://transformingexams.com/files/Hillier_2015_ascilite_fp.pdf)
10. Fluck, A., Pálsson, H., Coleman, M., Hillier, M., Schneider, D., Frankl, G., Uolia, K.: eExam symposium: design decisions and implementation experience. Presented at the IFIP World Conference on Computers in Education, 3-6 July, Dublin, Ireland (2017). Retrieved from: [http://transformingexams.com/files/Fluck\\_etal\\_2017.pdf](http://transformingexams.com/files/Fluck_etal_2017.pdf)
11. Hörnblad, P., Brenner, M.: Digital Exam. SUNET Incubator project, NORDUnet (2016). Retrieved from: <https://portal.nordu.net/display/Inkubator/Digital+Tentamen>
12. Dermo, J.: E-assessment and the student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology*, 40, 203–214 (2009)
13. Moge, N., Fluck, A.: Factors influencing student preference when comparing handwriting and typing for essay style examinations: Essay exams on computer. *British Journal of Educational Technology*, 46, 793-802 (2014)
14. Purcell, M., Paterson, J., Moge, N.: Exams: Comparing Handwritten Essays with those Composed on Keyboards (Final Report). Heslington: The HEA (2012)
15. Hillier, M.: The Very Idea of e-Exams: Student (Pre)conceptions. Proceedings of the ASCILITE conference, Dunedin, New Zealand, 77-88 (2014). Retrieved from: <http://ascilite.org/conferences/dunedin2014/files/fullpapers/91-Hillier.pdf>
16. Moge, N., Paterson, J., Burk, J., Purcell, M.: Typing compared with handwriting for essay examinations at university: letting the students choose. *Research in Learning Technology*, 18(1), 29-47 (2010)
17. Fluck, A.: eExaminations Strategic Project Final Report for Academic Senate (unpublished report). University of Tasmania, Australia (2011)
18. Rankin, A.: A comparability study on differences between scores of handwritten and typed responses on a large-scale writing assessment (unpublished PhD dissertation). University of Iowa, IA (2015)
19. Terzis, V., Economides, A.A.: Computer based assessment: Gender differences in perceptions and acceptance. *Computers and Human Behavior*, 27(6), 2108–2122 (2011)
20. Monash College Australia, Foundation Year (2016). Retrieved from: <https://web.archive.org/web/20160402024339/http://www.monashcollege.edu.au:80/courses/foundation-year>
21. Hillier, M., Lyon, N.: Student experiences with bring your own laptop e-Exams in pre-university college. In Passey, D., Bottino, R., Lewin, C., Sanchez, E. (Eds.) *Empowering Learners for Life in the Digital Age*, pp. 243-252. Heidelberg, Germany: Springer (2019)
22. Jamieson, S.: Likert scales: how to (ab)use them. *Medical Education*, 38(12), 1217–1218 (2004)
23. Mann, H.B., Whitney, D.R.: On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other. *Annals of Mathematics and Statistics*, 18(1), 50–60 (1947)
24. Freeman, G.H., Halton, J.H.: Note on an exact treatment of contingency tables, goodness of fit and other problems of significance. *Biometrika*, 38, 141–149 (1951)
25. Spearman, C.: The proof and measurement of association between two things. *American Journal of Psychology*, 15, 72–101 (1904)
26. Al Nadabi, Z.: Features of an online English language testing interface. Proceedings of the ASCILITE conference, 30 November - 3 December, Perth, Australia, 369–373 (2015)